7 QUANTILE TEST

The Quantile test was specifically developed to detect differences between the survey unit and the reference area that consist of a shift, δ', to higher values in only a fraction, ϵ, $0 < \epsilon < 1$, of the survey unit. It should be noted that, in general, this shift, δ', is not necessarily the same as the shift δ used for the WRS test. The Quantile test is only used in Scenario B. The Quantile test is performed after the WRS test, if the null hypothesis for that test has not been rejected. Using the Quantile test in tandem with the WRS test in Scenario B results in higher power to detect survey units that have not been adequately remediated than either test has by itself.

7.1 Introduction

The specific hypothesis tested by the Quantile test (see Johnson et al., 1987; EPA 230-R-94-004, 1994) is:

Null Hypothesis:
H$_0$: $\epsilon = 0$ or $\delta' \leq$ LBGR
versus
Alternative Hypothesis:
H$_a$: $\epsilon > 0$ and $\delta' >$ LBGR

Simply put, the null hypothesis is that there is no residual radioactivity above the LBGR in any part of the survey unit. The Quantile test is better at detecting situations in which only a portion, ϵ, of the survey unit contains excess residual radioactivity. The WRS test is better at detecting situations in which any excess residual radioactivity is uniform across the entire survey unit.

7.2 Applying the Quantile Test

For the Quantile test, the appropriate page in Table A.7 is selected, according to the value of $\alpha_Q = \alpha/2$.[1] Find the nearest values of *n*, the number of measurements from the survey unit, and *m*, the number of measurements from the reference area, that are tabulated. There are three numbers associated with each tabulated pair of *n* and *m* values, namely *r*, *k*, and α_Q.

The *m* measurements from the reference area and the *n* adjusted measurements from the survey unit are pooled and ranked in order of increasing size from 1 to *N*, where $N = m + n$. This is the same as steps (1) – (3) in Section 6.3 for the WRS test, and the same rankings can be used. If *k* or more of the *r* largest measurements in the combined ranked data set are from the survey unit, the null hypothesis is rejected For a survey unit that has failed the WRS test, it is not usually necessary to also perform the Quantile test. It is done here as an illustration of the method.

Table 7.1 reproduces the data in Table 6.5, but with two added columns showing the sorted ranks of the adjusted data, and the location associated with each rank, i.e., *R* for reference area and *S* for survey unit.

---

[1] Recall that since the Quantile test is performed in tandem with the WRS test, $\alpha_Q = \alpha_W \alpha/2$, so that the that the size of the two tests in tandem is approximately $\alpha = \alpha_Q + \alpha_W$.

Table 7.1 Quantile Test Under Scenario B for Class 2 Interior Drywall Survey Unit
(Measurements from the reference area and the survey unit are denoted by R and S, respectively)

	A	B	C	D	E	F	G
1	**Data**	**Area**	**Adjusted Data**	**Ranks**	**Survey Unit Ranks**	**Sorted Ranks**	**Location Associated With Sorted Rank**
2	47	R	47	18	—	1	R
3	28	R	28	1	—	2	R
4	36	R	36	6	—	3	R
5	37	R	37	7	—	4.5	R
6	39	R	39	9.5	—	4.5	S
7	45	R	45	13	—	6	R
8	43	R	43	11	—	7	R
9	34	R	34	3	—	8	S
10	32	R	32	2	—	9.5	R
11	35	R	35	4.5	—	9.5	R
12	39	R	39	9.5	—	11	R
13	51	R	51	21	—	13	R
14	209	S	67	24	24	13	S
15	197	S	55	23	23	13	S
16	188	S	46	16	16	16	S
17	191	S	49	19	19	16	S
18	193	S	51	21	21	16	S
19	187	S	45	13	13	18	R
20	188	S	46	16	16	19	S
21	180	S	38	8	8	21	R
22	193	S	51	21	21	21	S
23	188	S	46	16	16	21	S
24	187	S	45	13	13	23	S
25	177	S	35	4.5	4.5	24	S
26			**Sum =**	**300**	**194.5**		

On the second page of Table A.7, the closest entry to $n = m = 12$ is for $n = m = 10$. The values of $r = 7$, $k = 6$ and $\alpha_Q = 0.029$ are found. Thus, the null hypothesis is rejected if 6 of the 7 largest adjusted measurements come from the survey unit. From Table 7.1, we find that only 5 of the 7 largest adjusted measurements come from the survey unit.

The Quantile test as applied above yields only an approximate result. The values of n and m that were used are close to, but not equal to, the actual values. The value of α_Q will generally be different from that listed in the table. It is prudent to check a few other entries that are near the actual sample size. For $n = m = 15$, the values of $r = 5$, $k = 5$, and $\alpha_Q = 0.021$ are found. Thus, the null hypothesis is rejected if all of the 5 largest adjusted measurements come from the survey unit. From Table 7.1, we find that only 4 of the 5 largest adjusted measurements come from the survey unit. For $n = 15$, $m = 10$ the values of $r = 6$, $k = 6$, and $\alpha_Q = 0.028$ are found. From Table 7.1, we find that only 5 of the 6 largest adjusted measurements come from the survey unit. For $n = 10$, $m = 15$, we have $r = 6$, $k = 5$, and $\alpha_Q = 0.023$. Since 5 of the 6 largest adjusted measurements come from the survey unit, the null hypothesis would be rejected in this case. If the results are ambiguous, the methods of the next section can be used to fine tune the test for the sample sizes actually used..

The power of the Quantile test is more difficult to evaluate than that for the WRS test, since it depends on the two parameters δ' and ϵ. Setting specific values of these parameters in order to define a specific alternative for evaluating the power would often be speculative at best. It will be assumed that the power of the quantile test will be adequate for a range of values of δ' and ϵ when the sample size has been determined to assure adequate power for the WRS test.

If there are specific values of δ' and ϵ that are identified as being of concern during the historical site assessment, or prior surveys, power estimates can be obtained from tables in EPA 230-R-94-004 (1994), which list the power for various combinations of ϵ, δ'/σ, α_Q, m, n, r, and k. In those tables, only cases where $n = m$ are given. Since the power generally increases with sample size, upper and lower bounds for the power can be estimated when $m \neq n$ by consulting the entries for sample sizes both equal to the smaller or larger of these numbers.

7.3 Calculation of α_Q for the Quantile Test

The Quantile test, as applied in Section 7.2 gives only an approximate result, since the values used for n and m to find r and k were only approximately equal to the sample sizes actually used. Therefore, the actual value of α_Q will be different from that listed in the table. Fortunately, it is relatively easy to calculate the exact value of α_Q for specific values of n, m, k, and r. The number, k, out of the r largest measurements has a *hypergeometric* probability distribution when the null hypothesis is true. The probability that k or more of the r largest measurements are from the survey unit, when there is actually no residual radioactivity in the survey unit is:

$$\alpha_Q = \sum_{i=k}^{r} \frac{\binom{n}{i}\binom{m}{r-i}}{\binom{n+m}{r}} \tag{7-1}$$

The symbol $\binom{n}{i} = \frac{n!}{(n-i)!\ i!}$ is called a binomial coefficient. The symbol $n!$, called n factorial, is the product of the first n integers, $n! = n(n-1)(n-2)...(3)(2)(1)$. 0! is defined as equal to 1.

These calculations can be performed easily with any spreadsheet program that has the hypergeometric function built in. Otherwise, the following approximation, correct to almost three decimal places (Ling and Pratt, 1984), may be used:

$$1 - \alpha_Q = \Phi(z_{1-\alpha_Q}) \tag{7-2}$$

where

$$z_{1-\alpha_Q} = 2\,\frac{\sqrt{k(m-r+k)} - \sqrt{(n-k+1)(r-k+1)}}{\sqrt{(n+m-1)}}.$$

$\Phi(z)$ is the cumulative distribution function of a standard normal random variable tabulated in Table A.1. For the actual values of n and m, α_Q can be calculated for different combinations of r and k until a value sufficiently near the DQO is found.

Table 7.2 shows the calculations for the example data used in the previous section. For each of the possible values of r considered, i.e., 5, 6, and 7, the values of α_Q are calculated for each possible value of k, from 0 up to r. The value of α_Q closest to the desired value of $\alpha/2 = 0.025$ occurs for r = 5 and k = 5, where $\alpha_Q = 0.0186$. Using these values, and observing that only 4 of the 5 largest measurements are from the survey unit, the null hypothesis is not rejected. The survey unit passes the Quantile test, but it fails the WRS test (see Section 6.3). This may seem paradoxical, but is actually consistent with the patterns of residual radioactivity each test is designed to detect. This is discussed further in the next section.

Table 7.2 Example Calculation of α_Q for the Quantile Test

	A	B	C	D	E	F	G	H	I	J	K
1	*r* =	5			*r* =	6			*r* =	7	
2	*k* =	Prob	α		*k* =	Prob	α		*k* =	Prob	α
3	0	0.0186	1.0000		0	0.0069	1.0000		0	0.0023	1.0000
4	1	0.1398	0.9814		1	0.0706	0.9931		1	0.0320	0.9977
5	2	0.3416	0.8416		2	0.2427	0.9225		2	0.1510	0.9657
6	3	0.3416	0.5000		3	0.3596	0.6798		3	0.3146	0.8146
7	4	0.1398	0.1584		4	0.2427	0.3202		4	0.3146	0.5000
8	5	0.0186	0.0186		5	0.0706	0.0775		5	0.1510	0.1854
9					6	0.0069	0.0069		6	0.0320	0.0343
10									7	0.0023	0.0023
11		mean *k* =	2.50			mean *k* =	3.00			mean *k* =	3.50
12		std dev =	1.02			std dev =	1.08			std dev=	1.14

The spreadsheet formulas used for the example in Table 7.2 are shown in Table 7.3. In rows 1–12, only the formulas for columns A, B, and C are given, showing the calculations for $r = 5$, provided the values of n and m are defined in the spreadsheet. The formulas for the other columns are similar. One can get a feeling for the likelihood of the observed value of k by calculating the mean and standard deviation for k when the null hypothesis is true. For $n = m = 12$, and $r = 5$, the expected value of k is 2.5 ± 1.0. Thus, the observed value, $k = 4$, is 1.5 standard deviations above the mean. This is on the high end, but not quite high enough to reject the null hypothesis.

Table 7.3 Spreadsheet Formulas Used in Table 7.2

	A	B	C
1	*r* =	5	
2	*k* =	Prob	α
3	0	=HYPGEOMDIST(A3, *n, r, n+m*)	=1
4	1	=HYPGEOMDIST(A4, *n, r, n+m*)	=1-SUM(B$3:B3)
5	2	=HYPGEOMDIST(A5, *n, r, n+m*)	=1-SUM(B$3:B4)
6	3	=HYPGEOMDIST(A6, *n, r, n+m*)	=1-SUM(B$3:B5)
7	4	=HYPGEOMDIST(A7, *n, r, n+m*)	=1-SUM(B$3:B6)
8	5	=HYPGEOMDIST(A8, *n, r, n+m*)	=1-SUM(B$3:7)
9			
10			
11		mean *k* =	= (*n***r*) / (*m+n*)
12		std dev =	= sqrt(*m***n***r**(*m+n*-1) / (*m+n*)^2)

7.4 Modified Example for the Quantile Test

It was noted in the previous section that for the example data of Table 7.1 under Scenario B, the survey unit would fail the WRS test, yet pass the Quantile test. This occurred because the residual radioactivity is more or less uniformly distributed across the survey unit. If the example data is modified slightly, the result is different. Table 7.4 shows what would happen if 15 units of residual radioactivity are subtracted from the first six survey unit measurements and added to the last six survey unit measurements. The total residual radioactivity measured is unchanged. The analysis shows, however, that the sum of the adjusted survey unit measurement ranks is now 178. This is below the critical value of 184 from Section 6.3, and so the survey unit would pass the WRS test. Now, however, all of the highest ranked five measurements are from the survey unit, and so the survey unit would fail the Quantile test. This is because the spatial distribution of residual radioactivity is not uniform over the survey unit, but is concentrated at higher values in half of the survey unit. The purpose of using the two tests in tandem under Scenario B is to discover survey units with residual radioactivity in excess of the LBGR. There would be no value in using two tests if they always gave the same result.

Table 7.4 Quantile Test Under Scenario B: Modified Example Survey Unit
(Measurements from the reference area and the survey unit are denoted by R and S, respectively)

	A	B	C	D	E	F	G
1	**Data**	**Area**	**Adjusted Data**	**Ranks**	**Survey Unit Ranks**	**Sorted Ranks**	**Location Associated With Sorted Rank**
2	47	R	47	16	—	1	R
3	28	R	28	1	—	2	S
4	36	R	36	8.5	—	3	S
5	37	R	37	10	—	4	R
6	39	R	39	11.5	—	5.5	R
7	45	R	45	15	—	5.5	S
8	43	R	43	14	—	7	R
9	34	R	34	5.5	—	8.5	R
10	32	R	32	4	—	8.5	S
11	35	R	35	7	—	10	R
12	39	R	39	11.5	—	11.5	R
13	51	R	51	18	—	11.5	R
14	194	S	52	19	19	13	S
15	182	S	40	13	13	14	R
16	173	S	31	3	3	15	R
17	176	S	34	5.5	5.5	16	R
18	178	S	36	8.5	8.5	17	S
19	172	S	30	2	2	18	R
20	203	S	61	22.5	22.5	19	S
21	195	S	53	20	20	20	S
22	208	S	66	24	24	21	S
23	203	S	61	22.5	22.5	22.5	S
24	202	S	60	21	21	22.5	S
25	192	S	50	17	17	24	S
26			**Sum =**	**300**	**178**		